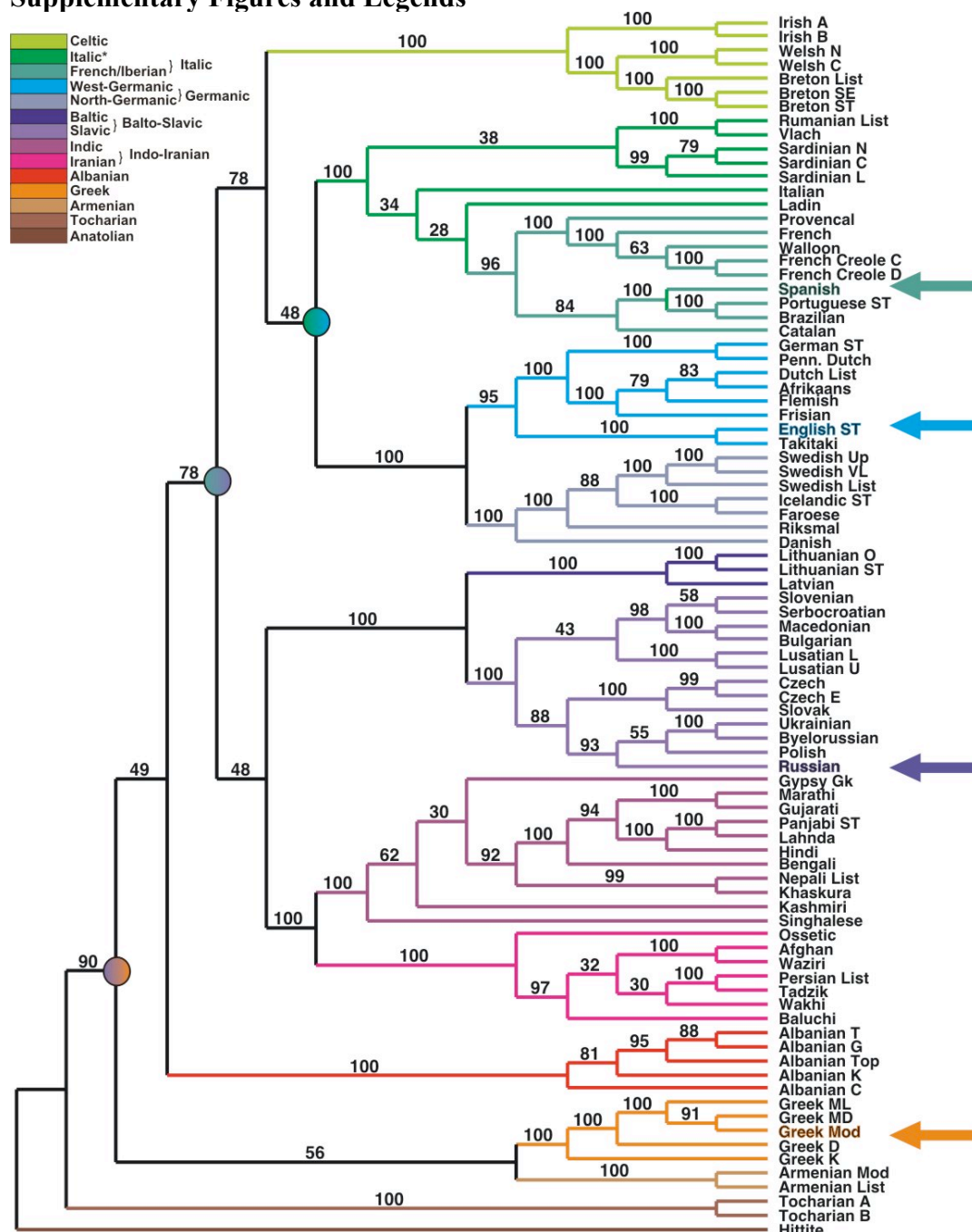
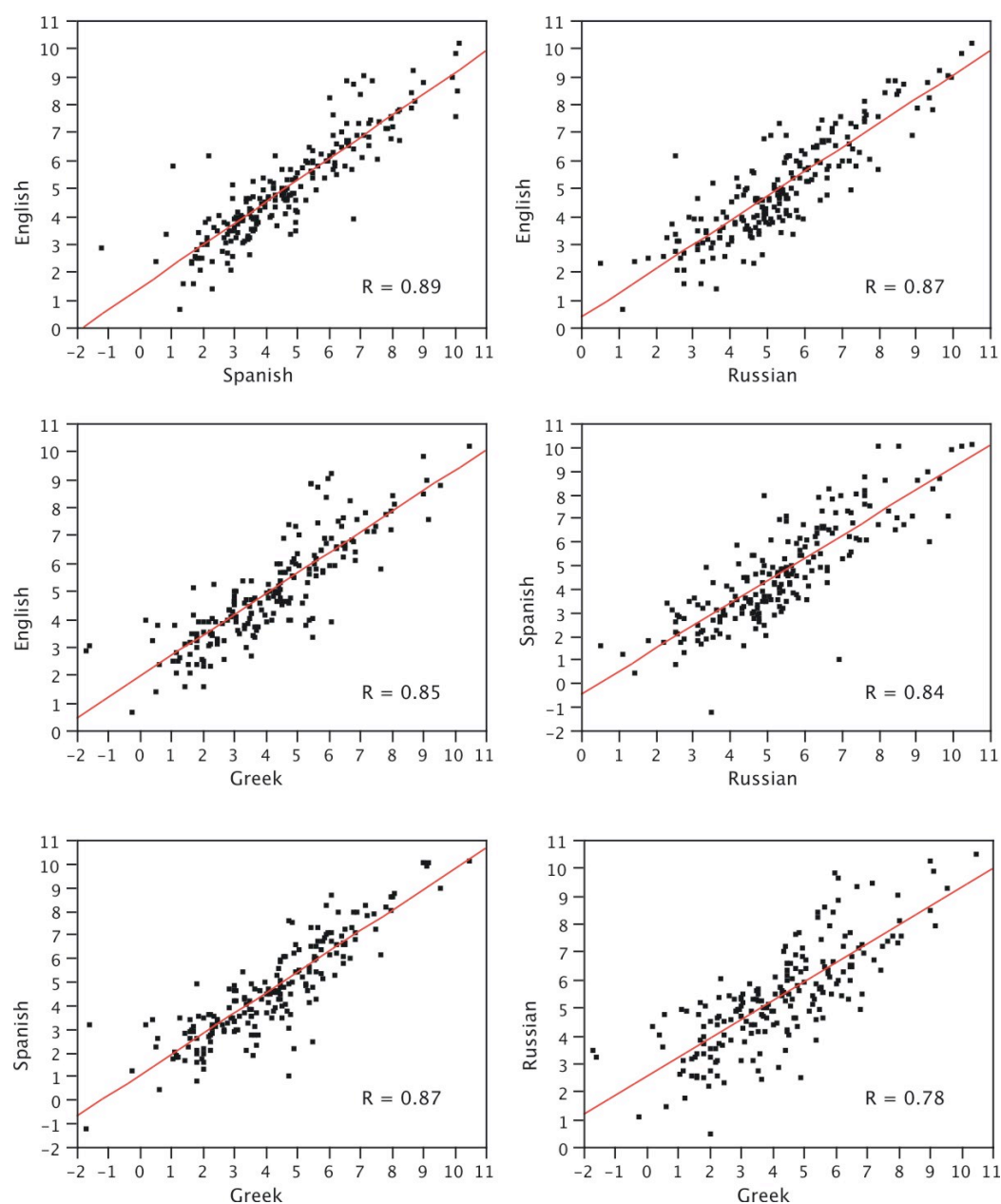


SUPPLEMENTARY INFORMATION

Supplementary Figures and Legends



Supplementary Figure S1 | Majority-rule consensus tree from the Bayesian MCMC sample of 750 Indo-European trees. Trees were produced using a two-state time reversible model of cognate gain and loss implemented in *BayesPhylogenies*^{30,33}. Values above each branch indicate uncertainty (posterior probability) in the tree as a percentage. The major accepted Indo-European language groupings are colour coded. These groupings are monophyletic and supported by high posterior probability values. Coloured arrows highlight the four languages from which we analysed corpus data. Coloured dots correspond to the most recent common ancestor of the corpora and show that comparisons among corpora sample deep nodes of the Indo-European phylogeny.



Supplementary Figure S2 | Pair-wise correlations between log-frequency of word use in English, Spanish, Russian and Greek. Values are taken from Supplementary Table S2.

Supplementary Tables

Supplementary Table S1 | Parallels between biological and linguistic evolution^{37, 38}

Biological Evolution	Language Evolution
Discrete heritable units – e.g. genetic code, morphology, behaviour	Discrete heritable units – e.g. lexicon, syntax, and phonology
Homology	Cognates
Mutation – e.g. Base-pair substitutions	Innovation – e.g. Sound changes
Drift	Drift
Natural selection	Social selection
Cladogenesis – e.g. allopatric speciation (geographic separation) and sympatric speciation (ecological/reproductive separation)	Lineage splits – e.g. geographical separation and social separation
Anagenesis	Change without split
Horizontal gene transfer – e.g. hybridisation	Borrowing
Plant Hybrids – e.g. wheat, strawberry	Language Creoles – e.g. Surinamese
Correlated genotypes/phenotypes – e.g. allometry, pleiotropy.	Correlated cultural terms – e.g. ‘five’ and ‘hand’.
Geographic clines	Dialects/Dialect chains
Fossils	Ancient Texts
Extinction	Language death

Similar parallels were first noted by Charles Darwin in *The Descent of Man*³⁹.

Supplementary Table S2 | Rate estimates and corpus frequency data across 200 basic vocabulary terms.

Meaning	Type	States	Mean Rate estimate (changes per 10,000 years)	S.D. Rate	English freq per million words	Spanish freq per million words	Russian freq per million words	Greek freq per million words
ALL	Adj.	15	2.02	0.03	2629.00	5477.21	8512.22	2874.09
AND	Conj.	25	3.39	0.03	26817.00	25203.09	36209.75	35683.41
ANIMAL	Noun	23	3.40	0.03	153.00	89.56	85.26	60.33
ASHES	Noun	21	2.65	0.03	13.00	5.74	8.88	6.99
AT	Prep.	28	5.69	0.04	4790.00	23610.29	4956.00	7955.55
BACK	Noun	26	4.29	0.03	212.00	48.38	297.72	34.05
BAD	Adj.	40	6.87	0.04	264.00	237.50	101.03	269.39
BARK (OF A TREE)	Noun	27	3.79	0.03	5.00	3.82	15.85	7.50
BECAUSE	Conj.	39	7.11	0.04	852.00	3855.15	777.33	365.24
BELLY	Noun	28	4.39	0.03	10.00	2.94	104.28	9.12
BIG	Adj.	24	3.41	0.03	338.00	478.09	1617.51	2040.45
BIRD	Noun	15	2.15	0.03	93.00	28.53	111.44	20.10
TO BITE	Verb	26	4.67	0.04	22.00	17.35	13.40	4.03
BLACK	Adj.	15	1.91	0.03	264.00	145.15	276.36	126.06
BLOOD	Noun	15	2.19	0.04	102.00	95.00	255.31	74.50
TO BLOW (WIND)	Verb	24	4.31	0.04	51.00	8.82	25.11	7.32
BONE	Noun	13	1.34	0.03	47.00	26.32	79.55	30.43
TO BREATHE	Verb	18	2.48	0.04	35.00	33.68	89.47	7.26
TO BURN (INTRANSITIVE)	Verb	20	2.94	0.03	53.00	33.53	18.03	64.37
CHILD (YOUNG)	Noun	36	6.31	0.03	710.00	735.88	557.01	500.84
CLOUD	Noun	24	3.25	0.03	37.00	19.71	45.52	22.16

COLD (WEATHER)	Adj.	22	3.58	0.03	103.00	104.85	156.77	19.88
TO COME	Verb	14	1.90	0.04	1512.00	1043.53	207.84	615.15
TO COUNT	Verb	27	4.07	0.03	50.00	853.97	395.23	426.37
TO CUT	Verb	23	3.36	0.03	184.00	70.00	33.38	79.38
DAY (NOT NIGHT)	Noun	10	1.04	0.03	940.00	1459.56	1281.93	827.10
TO DIE	Verb	8	0.82	0.04	220.00	348.53	63.47	85.11
TO DIG	Verb	27	3.86	0.03	28.00	2.21	12.37	6.16
DIRTY	Adj.	46	9.27	0.04	27.00	31.18	93.41	11.27
DOG	Noun	15	1.79	0.03	124.00	128.82	209.56	19.62
TO DRINK	Verb	12	1.49	0.03	75.00	71.76	173.80	47.46
DRY (SUBSTANCE)	Adj.	13	1.83	0.04	56.00	23.97	113.62	9.58
DULL (KNIFE)	Adj.	34	6.79	0.04	18.00	0.29	32.88	0.18
DUST	Noun	26	3.63	0.03	26.00	20.44	98.08	14.14
EAR	Noun	6	0.88	0.05	59.00	101.91	244.40	21.07
EARTH (SOIL)	Noun	23	3.52	0.03	97.00	137.65	622.47	110.93
TO EAT	Verb	15	1.77	0.03	144.00	270.29	1384.06	81.88
EGG	Noun	13	1.57	0.04	62.00	22.35	45.75	5.44
EYE	Noun	8	0.93	0.04	392.00	235.59	1308.72	141.69
TO FALL (DROP)	Verb	23	2.97	0.03	273.00	225.88	99.19	210.14
FAR	Adj.	19	3.14	0.04	288.00	90.88	396.38	84.16
FAT (SUBSTANCE)	Noun	31	4.89	0.03	21.00	14.71	34.22	8.64
FATHER	Pronoun	9	2.30	0.06	252.00	499.12	542.12	132.02
TO FEAR	Verb	25	4.06	0.03	53.00	39.12	335.93	116.92
FEATHER (LARGE)	Noun	15	2.35	0.04	13.00	18.68	45.94	13.94
FEW	Adj.	31	4.45	0.03	450.00	1169.85	250.25	940.23
TO FIGHT	Verb	29	4.21	0.03	108.00	35.29	63.59	33.35
FIRE	Noun	15	1.75	0.03	145.00	50.59	224.03	72.74
FISH	Noun	11	1.45	0.04	105.00	18.97	111.70	19.82
FIVE	Number	1	0.09	0.09	409.00	377.50	318.89	256.31
TO FLOAT	Verb	27	4.04	0.03	20.00	7.06	51.60	5.88
TO FLOW	Verb	27	3.40	0.03	26.00	8.24	40.89	6.14
FLOWER	Noun	18	2.29	0.03	76.00	109.56	129.13	17.17
TO FLY	Verb	21	2.83	0.03	90.00	53.38	55.32	91.16
FOG	Noun	32	4.97	0.04	11.00	5.15	77.37	6.03
FOOT	Noun	12	1.39	0.03	214.00	142.21	732.34	87.56
FOUR	Number	2	0.38	0.13	465.00	406.32	260.93	273.19
TO FREEZE	Verb	20	2.69	0.03	20.00	8.24	22.55	28.26
FRUIT	Noun	23	2.95	0.03	51.00	22.65	37.98	10.96
TO GIVE	Verb	7	0.76	0.04	1284.00	2657.21	590.85	1700.16
GOOD	Adj.	25	3.04	0.03	1276.00	2433.68	824.88	1368.02
GRASS	Noun	22	2.73	0.03	45.00	7.94	141.19	2.92
GREEN	Adj.	19	2.46	0.03	101.00	58.82	205.69	99.51
GUTS	Noun	35	6.89	0.04	5.00	5.44	24.50	4.21
HAIR	Noun	24	3.61	0.03	150.00	69.26	177.09	21.20
HAND	Noun	8	0.82	0.04	532.00	446.18	1772.56	254.92
HE	Pronoun	11	2.45	0.05	8469.00	1201.62	18972.22	389.45
HEAD	Noun	17	2.32	0.03	402.00	195.29	1112.87	75.09
TO HEAR	Verb	15	2.35	0.04	367.00	471.32	346.07	405.63
HEART	Noun	13	1.69	0.03	152.00	106.47	280.73	72.39
HEAVY	Adj.	20	2.89	0.03	105.00	32.65	246.08	91.11
HERE	Adv.	24	3.25	0.03	699.00	1438.97	927.16	676.66

TO HIT	Verb	40	5.51	0.03	107.00	44.41	106.39	103.54
HOLD (IN HAND)	Verb	29	4.48	0.03	481.00	71.47	266.87	262.12
HOW	Adv.	3	0.41	0.08	1016.00	1206.47	7165.64	427.02
TO HUNT (GAME)	Verb	23	3.33	0.03	15.00	16.03	15.85	34.62
HUSBAND	Noun	28	4.74	0.03	123.00	126.18	226.56	191.30
I	Pronoun	1	0.09	0.09	10241.00	5990.29	15227.54	430.75
ICE	Noun	20	2.58	0.03	41.00	16.32	62.09	41.06
IF	Conj.	27	4.32	0.03	2369.00	3533.97	1979.70	2504.48
IN	Prep.	12	1.78	0.04	18214.00	22657.50	27965.58	7955.55
TO KILL	Verb	28	3.50	0.03	157.00	163.97	68.79	83.46
KNOW (FACTS)	Verb	13	2.11	0.04	1882.00	2906.47	1985.18	532.19
LAKE	Noun	25	3.50	0.03	51.00	19.71	72.66	21.00
TO LAUGH	Verb	13	1.61	0.03	98.00	106.76	152.13	43.69
LEAF	Noun	19	2.43	0.03	51.00	40.15	135.37	55.29
LEFT (HAND)	Adj.	30	4.41	0.03	95.00	191.62	130.47	173.70
LEG	Noun	25	3.73	0.03	118.00	70.88	732.34	87.56
TO LIE (ON SIDE)	Verb	23	3.56	0.03	187.00	35.44	431.75	10.55
TO LIVE	Verb	10	1.10	0.04	329.00	668.24	781.35	264.40
LIVER	Noun	17	2.66	0.04	17.00	6.03	22.93	3.09
LONG	Adj.	12	1.22	0.03	573.00	237.94	368.04	142.79
LOUSE	Noun	12	1.74	0.04	2.00	3.53	2.99	0.77
MAN (MALE)	Noun	18	3.38	0.04	1003.00	964.12	239.15	347.29
MANY	Adj.	32	5.60	0.04	902.00	2866.32	137.93	865.82
MEAT (FLESH)	Noun	15	1.62	0.03	38.00	78.53	83.57	27.54
MOTHER	Pronoun	10	2.36	0.07	295.00	447.94	423.52	97.89
MOUNTAIN	Noun	20	2.75	0.03	68.00	34.56	191.37	33.61
MOUTH	Noun	19	2.78	0.03	99.00	102.21	105.47	45.12
NAME	Noun	4	0.47	0.06	326.00	280.74	348.41	214.85
NARROW	Adj.	19	2.89	0.04	54.00	20.59	104.40	114.44
NEAR	Adj.	35	4.86	0.03	44.00	138.24	172.00	188.69
NECK	Noun	21	3.57	0.04	60.00	32.50	140.38	17.60
NEW	Adj.	5	0.60	0.05	1154.00	589.85	850.37	151.47
NIGHT	Noun	6	0.76	0.05	393.00	391.32	567.12	85.39
NOSE	Noun	14	1.49	0.03	48.00	23.97	225.29	13.88
NOT	Adv.	7	0.82	0.04	7995.00	20062.35	20408.27	8820.30
OLD	Adj.	18	2.53	0.03	648.00	215.59	508.01	317.08
ONE	Number	2	0.43	0.14	1962.00	22858.68	2872.91	9484.09
OTHER	Adj.	18	2.66	0.03	1336.00	3098.38	1577.31	2806.21
PERSON	Noun	28	4.81	0.04	290.00	829.26	2884.62	338.61
TO PLAY	Verb	24	3.18	0.03	386.00	197.50	231.30	495.14
TO PULL	Verb	28	4.53	0.03	140.00	88.38	83.11	56.17
TO PUSH	Verb	40	8.06	0.04	107.00	28.53	22.97	20.28
TO RAIN	Verb	19	2.39	0.03	14.00	32.79	128.86	3.42
RED	Adj.	14	2.26	0.04	126.00	96.03	316.56	68.69
RIGHT (CORRECT)	Adj.	25	3.91	0.03	52.00	62.65	48.12	213.45
RIGHT (HAND)	Adj.	18	2.61	0.03	1000.00	456.03	162.82	97.16
RIVER	Noun	23	3.69	0.03	114.00	109.71	196.39	38.23
ROAD	Noun	31	4.86	0.03	313.00	172.50	448.56	338.94
ROOT	Noun	13	1.75	0.04	45.00	41.91	60.98	28.00
ROPE	Noun	35	6.09	0.03	22.00	27.35	55.24	11.22
ROTTEN (LOG)	Adj.	32	4.76	0.03	8.00	6.76	15.73	3.11

RUB	Verb	27	3.57	0.03	23.00	8.38	12.90	9.32
SALT	Noun	10	1.01	0.03	33.00	17.06	41.54	9.93
SAND	Noun	27	4.26	0.03	34.00	38.38	82.19	9.45
TO SAY	Verb	26	3.64	0.03	3344.00	6301.62	1976.10	3133.88
SCRATCH (ITCH)	Verb	23	4.44	0.04	12.00	6.18	6.01	3.33
SEA (OCEAN)	Noun	15	2.53	0.04	139.00	98.82	190.23	81.53
TO SEE	Verb	20	2.63	0.03	1920.00	2841.76	1047.91	997.67
SEED	Noun	14	2.56	0.04	32.00	11.76	22.43	9.32
TO SEW	Verb	13	1.71	0.04	8.00	17.79	12.75	4.93
SHARP (KNIFE)	Adj.	24	3.76	0.03	44.00	14.12	117.91	1.75
SHORT	Adj.	24	3.39	0.03	198.00	105.44	198.04	202.31
TO SING	Verb	20	2.54	0.03	63.00	147.94	136.86	53.01
TO SIT	Verb	11	1.25	0.03	300.00	117.21	735.94	97.54
SKIN (OF PERSON)	Noun	17	2.73	0.04	75.00	51.47	130.16	16.44
SKY	Noun	18	2.52	0.03	56.00	65.44	270.35	32.32
TO SLEEP	Verb	15	2.01	0.03	68.00	142.21	240.99	33.56
SMALL	Adj.	26	3.70	0.03	518.00	318.68	396.34	644.06
TO SMELL (PERCEIVE ODOR)	Verb	36	6.48	0.04	25.00	31.32	9.99	11.79
SMOKE	Noun	11	1.70	0.04	29.00	15.88	116.88	38.21
SMOOTH	Adj.	35	6.23	0.04	30.00	6.47	28.98	35.71
SNAKE	Noun	20	3.11	0.04	12.00	7.21	34.22	7.65
SNOW	Noun	11	1.62	0.04	31.00	21.91	213.12	9.87
SOME	Adj.	26	2.57	0.02	1712.00	1543.82	654.82	276.76
TO SPIT	Verb	17	2.04	0.03	11.00	9.85	25.50	4.63
TO SPLIT	Verb	31	5.58	0.04	32.00	140.59	29.02	6.09
TO SQUEEZE	Verb	42	7.83	0.04	21.00	25.29	25.53	0.20
TO STAB (OR STICK)	Verb	40	7.84	0.04	11.00	1.62	4.21	1.84
TO STAND	Verb	16	2.02	0.03	326.00	2.79	1003.65	111.22
STAR	Noun	7	0.72	0.04	100.00	73.38	129.97	28.08
STICK (OF WOOD)	Noun	38	8.02	0.04	26.00	29.71	57.50	1.47
STONE	Noun	21	2.73	0.03	112.00	59.85	204.96	28.52
STRAIGHT	Adj.	27	4.08	0.03	32.00	19.12	114.96	5.11
TO SUCK	Verb	19	2.41	0.03	16.00	13.38	12.63	4.74
SUN	Noun	8	0.98	0.04	115.00	115.74	244.59	58.45
TO SWELL	Verb	29	4.54	0.03	10.00	5.15	1.65	7.30
TO SWIM	Verb	22	2.73	0.03	25.00	13.97	41.77	6.09
TAIL	Noun	28	4.16	0.03	32.00	36.18	94.67	22.82
THAT	Adj.	18	1.88	0.03	3844.00	401.91	11307.07	777.73
THERE	Adv.	22	2.94	0.03	746.00	499.85	1334.30	509.70
THEY	Pronoun	14	2.75	0.05	6081.00	858.82	5654.32	274.57
THICK	Adj.	29	5.41	0.04	51.00	36.18	163.08	6.34
THIN	Adj.	25	2.99	0.03	56.00	12.06	137.55	234.87
TO THINK	Verb	36	5.34	0.03	1520.00	710.00	929.58	139.87
THIS	Adj.	15	2.18	0.04	4623.00	5582.50	3457.94	3013.62
THOU	Pronoun	6	0.64	0.04	6954.00	691.47	4628.01	228.97
THREE	Number	1	0.09	0.09	800.00	734.41	669.22	685.03
TO THROW	Verb	42	7.52	0.03	115.00	88.38	77.22	115.52
TO TIE	Verb	21	2.68	0.03	42.00	22.06	11.18	36.94
TONGUE	Noun	4	0.49	0.06	28.00	116.47	289.99	241.56
TOOTH (FRONT)	Noun	7	1.18	0.05	54.00	26.03	201.02	16.10

TREE	Noun	26	3.58	0.03	147.00	67.94	214.34	19.59
TO TURN (VEER)	Verb	34	8.18	0.05	465.00	8.82	12.37	131.05
TWO	Number	1	0.09	0.09	1563.00	1412.94	1366.57	1783.49
TO VOMIT	Verb	32	5.02	0.03	4.00	9.56	36.90	1.65
TO WALK	Verb	36	5.14	0.03	215.00	232.65	353.16	25.60
WARM (WEATHER)	Adj.	16	2.35	0.04	70.00	32.35	143.52	18.93
TO WASH	Verb	24	3.00	0.03	49.00	38.24	20.83	8.56
WATER	Adj.	8	0.86	0.04	372.00	241.62	598.74	156.69
WE	Pronoun	2	0.37	0.12	4202.00	1097.06	4731.37	376.82
WET	Adj.	28	4.45	0.03	37.00	10.29	108.45	40.96
WHAT	Adv.	5	0.69	0.06	2493.00	3880.29	12640.15	1311.60
WHEN	Adv.	10	1.08	0.03	431.00	1899.71	2268.20	124.64
WHERE	Adv.	7	0.70	0.04	628.00	884.85	1466.37	920.19
WHITE	Adj.	17	2.22	0.03	207.00	121.62	481.02	86.28
WHO	Pronoun	1	0.09	0.09	2055.00	452.35	2154.42	626.06
WIDE	Adj.	20	2.89	0.03	165.00	18.97	180.92	5.29
WIFE	Noun	31	4.12	0.03	190.00	62.50	445.30	58.44
WIND (BREEZE)	Noun	16	1.78	0.03	85.00	47.79	204.28	31.35
WING	Noun	22	2.66	0.03	53.00	23.97	74.73	1.21
WIPE	Verb	39	6.92	0.03	24.00	30.29	29.40	4.94
WITH (ACCOMPANYING)	Prep.	19	3.25	0.04	6575.00	7876.76	11130.94	13745.15
WOMAN	Noun	25	2.75	0.02	631.00	591.62	564.13	304.61
WOODS	Noun	27	4.15	0.03	67.00	43.53	319.35	67.43
WORM	Noun	19	2.16	0.03	12.00	5.88	14.01	2.85
YE	Pronoun	13	1.32	0.03	6984.00	1549.41	3896.62	228.97
YEAR	Noun	21	2.78	0.03	1639.00	1985.15	2039.27	111.11
YELLOW	Adj.	18	2.12	0.03	41.00	40.29	151.94	41.02

For each meaning the table shows: part of speech; observed number of states; mean estimated rate of replacement per 10,000 years from across the MCMC sample distribution; standard deviation of rate estimates across the MCMC sample distribution; and the frequency of word use per million words in English ², Spanish ³, Russian ⁴ and Greek ⁵. Word frequencies were based on the lemma frequencies of each meaning. For example, for the meaning ‘push’ in English, we include ‘push’, ‘pushes’, ‘pushing’, and ‘pushed’.

Supplementary Discussion

Grammaticalization

Grammaticalization describes the process of language change in which a word changes its part of speech over time⁴⁰. This does not affect the estimates of the rates of lexical replacement associated with meanings, reported in Figure 1, because the estimates of rates of lexical replacement are associated with specific meanings, not with the rates of particular words. If a word within a given language changes its part of speech and is replaced with another word for that meaning, then the rate of lexical replacement for that meaning correctly includes the change. Equally, if a word from one meaning changes its part of speech and thereby moves into another meaning category (such as the example of ‘like’ moving from being an adjective to a preposition) the estimate of the rate of lexical replacement for the latter meaning category (i.e., ‘like’ as a preposition) will correctly include that change.

The Indo-European language database³¹ and the corpus data identify the part of speech that a particular word represents. Thus, for example, both would distinguish between using ‘hit’ as a noun (as in ‘she was a hit with the press’) and ‘hit’ as a verb (the Indo-European database only includes the latter). Both sources also, for example, distinguish between the two meanings associated with the adjective ‘right’. Combined with the above description of how rates are estimated, this means both the rates of lexical replacement and estimates of the frequencies of use avoid problems associated with grammaticalization. The finding that the rank order of the effects of parts of speech was identical in the four languages (Figure 3), despite the fact that grammaticalization would have been taking place in them independently, reinforces this conclusion.

Grammaticalization could have an effect if there had been a recent, specific and law-like trend of grammaticalization across all Indo-European languages: for example, if the words used in the various Indo-European languages for the same noun all changed to being the same verb. The estimated rate of lexical replacement for this verb meaning would then reflect the historical rate of lexical replacement for the noun forms. We know of no evidence for such widespread and recent changes. But even if such trends did exist they would not render invalid the findings that rates of lexical replacement differ for different parts of speech, they would simply qualify what could be said about particular parts of speech.

Calculating rates of lexical replacement for meanings with one cognate class

Meanings with a single cognate class have an observed rate of lexical replacement equal to zero, and this is also their maximum likelihood statistical estimate. A rate of zero yields an infinite half-life and so it is desirable to try to estimate some positive value for these rates. This can be achieved by asking what value of a rate of evolution is still compatible with observing no changes over the evolutionary time period represented by the phylogenetic tree of the Indo-European languages (Supplementary Figure S1). The Markov chain Monte Carlo techniques we describe in the Methods find the posterior distribution of rate coefficients for each of the 200 meanings we investigated. For meanings with a single cognate class, the mean of the posterior distribution was 0.09 replacements per 10,000 years and this yields a linguistic half-life of 76,530 years. This may seem an unreasonably long half-life but it should be borne in mind that the phylogenetic tree for the Indo-European languages represents

130,000 years of linguistic evolution if it is assumed that Indo-European is 8,700 years old. During this time there has not been any lexical replacement for the meanings with a single cognate class.

Supplementary Notes

37. Pagel, M. in *Time Depth in Historical Linguistics* (eds. Renfrew, C., McMahon, A. & Trask, L.) 189-207 (MacDonald Institute of Archaeological Research, Cambridge, 2000).
38. Atkinson, Q. D. & Gray, R. D. Curious parallels and curious connections--phylogenetic thinking in biology and historical linguistics. *Syst Biol* **54**, 513-26 (2005).
39. Darwin, C. *The Descent of Man* (John Murray, London, 1871).
40. Hopper, P. J. & Traugott, E. C. *Grammaticalization* (Cambridge University Press, Cambridge [England] ; New York, NY, USA, 1993).